



**E-Infrastructures
H2020-EINFRA-2016-2017**

EINFRA-11-2016: Support to the next implementation phase of Pan-European High Performance Computing Infrastructure and Services (PRACE)

PRACE-5IP

PRACE Fifth Implementation Phase Project

Grant Agreement Number: EINFRA-730913

D6.6

Final Report on Collaborations with e-Infrastructures and CoEs

Final

Version: 1.0
Author(s): Cedric Jourdain, CINES; Fabio Affinito, CINECA; Huub Stoffers, SURFSara, Miroslav Puskaric, HLRS, Joni Virtanen, CSC; Pedro Alberto, UC-LCA, Ralph Niederberger, JUELICH
Date: 26/04/2019

Project and Deliverable Information Sheet

| | | |
|---|--|--|
| PRACE Project | Project Ref. №: EINFRA-730913 | |
| | Project Title: PRACE Fifth Implementation Phase Project | |
| | Project Web Site: http://www.prace-project.eu | |
| | Deliverable ID: < D6.6> | |
| | Deliverable Nature: <Report> | |
| | Dissemination Level: PU * | Contractual Date of Delivery: 30/04/2019 |
| | | Actual Date of Delivery: 30/04/2019 |
| EC Project Officer: Leonardo Flores Añover | | |

* - 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 2005/444/EC.

Document Control Sheet

| | | |
|-------------------|--|--|
| Document | Title: Final Report on Collaborations with e-Infrastructures and CoEs | |
| | ID: D6.6 | |
| | Version: <1.0> | Status: <i>Final</i> |
| | Available at: http://www.prace-project.eu | |
| | Software Tool: Microsoft Word 2013 | |
| | File(s): PRACE-5IP-D6.6-v1.0 | |
| Authorship | Written by: | Cedric Jourdain, CINES; Fabio Affinito, CINECA; Huub Stoffers, SURFSara, Miroslav Puskaric, HLRS, Joni Virtanen, CSC; Pedro Alberto, UC-LCA, Ralph Niederberger, JUELICH |
| | Contributors: | |
| | Reviewed by: | Norbert Meyer, PSNC Veronica Teodor, JUELICH |
| | Approved by: | MB/TB |

Document Status Sheet

| Version | Date | Status | Comments |
|----------------|---------------|-------------------------|---|
| 0.1 | 15/March/2019 | Skeleton | Skeleton and Initial draft |
| 0.2 | 19/March/2019 | First integrated draft | Add references, abbreviations and minor corrections |
| 0.3 | 26/March/2019 | Second integrated draft | minor restructuring, addition of transitions, |

| | | | |
|-----|---------------|---------------|--|
| | | | addition of sections related to CoEs |
| 0.4 | 29/March/2019 | Draft | Addition of introduction and conclusion and minor correction |
| 0.5 | 01/April/2019 | Revised draft | Version submitted for internal review |
| 0.6 | 17/April/2019 | Semi-final | Integration of updates based on reviewer feedback |
| 1.0 | 26/April/2019 | Final | Addressing comments and suggestions from MB/TB reviewer |

Document Keywords

| | |
|------------------|---|
| Keywords: | PRACE, HPC, Research Infrastructure, Collaboration, Cooperation, e-Infrastructures, data pilots, Centres of Excellence, FET projects, Data management, Security, Training |
|------------------|---|

Disclaimer

This deliverable has been prepared by the responsible Work Package of the Project in accordance with the Consortium Agreement and the Grant Agreement n° EINFRA-730913. It solely reflects the opinion of the parties to such agreements on a collective basis in the context of the Project and to the extent foreseen in such agreements. Please note that even though all participants to the Project are members of PRACE AISBL, this deliverable has not been approved by the Council of PRACE AISBL and therefore does not emanate from it nor should it be considered to reflect PRACE AISBL's individual opinion.

Copyright notices

© 2019 PRACE Consortium Partners. All rights reserved. This document is a project document of the PRACE project. All contents are reserved by default and may not be disclosed to third parties without the written consent of the PRACE partners, except as mandated by the European Commission contract EINFRA-730913 for reviewing and dissemination purposes.

All trademarks and other rights on third party products mentioned in this document are acknowledged as own by the respective holders.

Table of Contents

| | |
|--|-----------|
| Document Control Sheet..... | i |
| Document Status Sheet | i |
| Document Keywords | iii |
| List of Figures | v |
| List of Tables..... | v |
| References and Applicable Documents | v |
| List of Acronyms and Abbreviations..... | vii |
| List of Project Partner Acronyms..... | viii |
| Executive Summary | 10 |
| 1 Introduction..... | 11 |
| 2 Collaboration with other e-Infrastructures..... | 12 |
| 2.1 GÉANT | 12 |
| 2.2 EUDAT | 13 |
| 2.2.1 <i>Leveraging EUDAT services.....</i> | <i>13</i> |
| 2.2.2 <i>DECI-13 and DECI-14 Pilots</i> | <i>15</i> |
| 2.2.2.1 <i>CHARTERED2 Data Pilot.....</i> | <i>16</i> |
| 2.2.2.1.1 <i>Project and actors.....</i> | <i>16</i> |
| 2.2.2.1.2 <i>Current status.....</i> | <i>17</i> |
| 2.2.2.2 <i>SubGridEoR Data Pilot</i> | <i>18</i> |
| 2.2.2.3 <i>Analysis of the outcomes</i> | <i>18</i> |
| 2.2.3 <i>Training activities</i> | <i>19</i> |
| 2.3 EGI..... | 19 |
| 2.4 e-InfraCentral | 20 |
| 2.5 Human Brain Project | 20 |
| 2.6 XSEDE / RIST | 21 |
| 2.7 Collaboration with CERN..... | 21 |
| 2.8 Collaboration to support Globus based services | 23 |
| 3 Collaboration with CoEs | 25 |
| 3.1 PRACE-EXDCI-CoEs meeting in Brühl, Germany..... | 26 |
| 3.2 BioExcel Centre of Excellence | 27 |
| 3.3 Energy oriented Centre of Excellence | 28 |
| 3.4 Performance Optimisation and Productivity Centre of Excellence..... | 28 |

| | | |
|----------|---|-----------|
| 4 | Collaboration on the security domain | 29 |
| 4.1 | Collaboration with WISE | 29 |
| 4.2 | Federated Authentication and Authorisation | 29 |
| 4.3 | Trusted Introducer service | 30 |
| 5 | Conclusion | 32 |

List of Figures

| | |
|---|----|
| Figure 1: PerfSONAR Dashboard [27] | 13 |
| Figure 2: EUDAT Data Domain modelled on the ANDS [28] Data Curation Continuum | 14 |
| Figure 3: Data flow diagram for CHARTERED2 data pilot..... | 17 |
| Figure 4: PRACE-CoEs-FET HPC-EXDCI Workshop – October 2018 | 27 |
| Figure 5: Team info of the PRACE CSIRT [45]..... | 31 |

List of Tables

| | |
|---|----|
| Table 1: DECI-13 call data pilots..... | 15 |
| Table 2: DECI-14 call data pilots..... | 16 |
| Table 3: Proposed activities for future collaboration with CERN | 23 |
| Table 4: Status of upgrade to GCT among PRACE partners | 24 |
| Table 5: List of Centres of Excellence | 25 |

References and Applicable Documents

- [1] <http://www.prace-project.eu/>
- [2] EGI <https://www.egi.eu/>
- [3] EUDAT project <http://www.eudat.eu/>
- [4] GEANT Project <http://www.geant.org/>
- [5] Human Brain Project <https://www.humanbrainproject.eu/>
- [6] Centres of Excellence <http://ec.europa.eu/programmes/horizon2020/en/news/eight-new-centres-excellence-computing-applications>
- [7] Energy Oriented CoE for computing applications <http://www.eocoe.eu/>
- [8] POP CoE in computing applications <https://pop-coe.eu/>
- [9] BioExcel CoE for Computational Biomolecular Research <https://bioexcel.eu/>
- [10] CoeGSS <http://coegss.eu/>
- [11] HiDALGO <https://hidalgo-project.eu/>
- [12] EXCELLERAT <https://www.excellerat.eu/wp/>
- [13] XSEDE <https://www.xsede.org/>
- [14] RIST <http://www.rist.or.jp/ehome.html>
- [15] EUDAT service descriptions <http://www.eudat.eu/services>
- [16] Data Management Plan <https://www.eudat.eu/d22-data-management-plan>

- [17] eduGAIN <http://services.geant.net/edugain/Pages/Home.aspx>
- [18] AARC project <https://aarc-project.eu/>
- [19] WISE <https://wise-community.org/>
- [20] EOSC <https://ec.europa.eu/digital-single-market/en/european-open-science-cloud>
- [21] European research area https://ec.europa.eu/info/research-and-innovation/strategy/era_en
- [22] PRACE-4IP D6.6 http://www.prace-ri.eu/IMG/pdf/D6.6_4ip.pdf
- [23] PRACE-4IP D6.5 http://www.prace-ri.eu/IMG/pdf/D6.5_4ip.pdf
- [24] perSONAR <https://www.perfsonar.net/>
- [25] MaDDash http://docs.perfsonar.net/maddash_intro.html
- [26] ESMOND <https://github.com/perfsonar/esmond>
- [27] PerfSONAR Dashboard:
<https://maddash.prace.fz-juelich.de/maddash-webui/index.cgi?dashboard=Dashboard%201>
- [28] Australian National Data Service organization <http://www.and.s.org.au/>
- [29] D7.3 EUDAT/PRACE Joint Call for Proposal <https://eudat.eu/d73-eudatprace-joint-call-for-proposals>
- [30] EUDAT data management training <https://www.eudat.eu/training/research-data-management>
- [31] European E-Infrastructure Services Gateway
<https://cordis.europa.eu/project/rcn/206232/factsheet/en>
- [32] E-InfraCentral: <http://einfracentral.eu/>
- [33] EOSC-hub <https://www.eosc-hub.eu/>
- [34] "Grid Community Forum," <https://gridcf.org/>
- [35] Scheiner, " RFC 281 - Upgrade to Grid Community Toolkit 6.2," 13 December 2018
<https://prace-wiki.fz-juelich.de/bin/view/PRACE/Operations/RFC281>
- [36] PRACE/BioExcel Spring School 2017 <https://www.pdc.kth.se/publications/pdc-newsletter-articles/2017-no-1/prace-bioexcel-spring-school-2017-1.744815>
- [37] PRACE/BioExcel Seasonal School 2019 <https://bioexcel.eu/events/bioexcel-prace-seasonal-school-2019-sweden-hpc-for-life-sciences/>
- [38] Scientific Applications towards Exascale workshop <https://events.prace-ri.eu/event/624/overview>
- [39] Laboratoire LI-PaRAD <http://www.liparad.uvsq.fr/>
- [40] Modular Assembly Quality Analyzer and Optimizer <http://www.maqao.org/>
- [41] Trust Framework for Security Collaboration among Infrastructures version 1
<https://pos.sissa.it/179/011/pdf>
- [42] Trust Framework for Security Collaboration among Infrastructures version 2 <https://wise-community.org/wp-content/uploads/2017/05/WISE-SCI-V2.0.pdf>
- [43] AARC Blueprint Architecture <https://aarc-project.eu/wp-content/uploads/2017/04/AARC-BPA-2017.pdf>
- [44] TF-CSIRT Trusted Introducer: <https://www.trusted-introducer.org/index.html>
- [45] Trusted Introducer team <https://www.trusted-introducer.org/directory/teams/prace-csirt.html>
- [46] ChEESE <https://cheese-coe.eu/>
- [47] FocusCoE <https://www.focus-coe.eu/>

List of Acronyms and Abbreviations

| | |
|-------|--|
| aisbl | Association International Sans But Lucratif (legal form of the PRACE-RI) |
| AARC | Authentication and Authorisation for Research and Collaboration |
| AEGIS | AARC ENGAGEMENT Group for Infrastructure |
| AEGIS | Antihydrogen Experiment: Gravity, Interferometry, Spectroscopy |
| BPA | Blueprint Architecture |
| CoE | Center of Excellence |
| CSA | Coordination and Support Actions |
| CSIRT | Computer Security Incident Response Team |
| DECI | DEISA Extreme Computing Initiative |
| DEISA | Distributed European Infrastructure for Supercomputing Applications EU project by leading national HPC centres |
| DMP | Data Management Plan |
| EC | European Commission |
| EDI | European Data Infrastructure |
| EGI | European Grid Infrastructure |
| EOSC | European Open Sciences Cloud |
| ESFRI | European Strategy Forum on Research Infrastructures |
| EUDAT | European Collaborative Data Infrastructure |
| FET | Future and Emerging Technologies |
| FENIX | Fusion Engineering International Experiments |
| GDPR | General Data Protection Regulation |
| GÉANT | Collaboration between National Research and Education Networks to build a multi-gigabit pan-European network. |
| GHz | Giga (= 10^9) Hertz, frequency = 10^9 periods or clock cycles per second |
| GPU | Graphic Processing Unit |
| GCT | Grid Community Toolkit |
| GSI | Grid Security Infrastructure |
| HEP | High Energy Physics |
| HET | High Performance Computing in Europe Taskforce. Taskforce by representatives from European HPC community to shape the European HPC Research Infrastructure. Produced the scientific case and valuable groundwork for the PRACE project. |
| HPC | High Performance Computing; Computing at a high performance level at any given time; often used synonym with Supercomputing |
| Idp | Identity Provider |
| ISC | International Supercomputing Conference; European equivalent to the US based SCxx conference. Held annually in Germany. |
| LHC | Large Hadron Collider |
| MB | Management Board (highest decision making body of the project) |
| MDVPN | Multi-Domain Virtual Private Network |
| MOOC | Massively open online Course |
| MoU | Memorandum of Understanding. |
| NAMD | Nanoscale Molecular Dynamics software |
| PA | Preparatory Access (to PRACE resources) |
| PTC | PRACE Training Centres |

| | |
|---------|--|
| PRACE | Partnership for Advanced Computing in Europe; Project Acronym |
| PRACE 2 | The upcoming next phase of the PRACE Research Infrastructure following the initial five year period. |
| RI | Research Infrastructure |
| SIM | Security Incident Management |
| SKA | Square Kilometre Array |
| SPS | Super Proton Synchrotron |
| TB | Technical Board (group of Work Package leaders) |
| TB | Tera (= $2^{40} \sim 10^{12}$) Bytes (= 8 bits), also TByte |
| TCO | Total Cost of Ownership. Includes recurring costs (e.g. personnel, power, cooling, maintenance) in addition to the purchase cost. |
| TI | Trusted Introducer |
| Tier-0 | Denotes the apex of a conceptual pyramid of HPC systems. In this context the Supercomputing Research Infrastructure would host the Tier-0 systems; national or topical HPC centres would constitute Tier-1 |
| VMD | Visual Molecular Dynamics, molecular graphics software |

List of Project Partner Acronyms

| | |
|------------------|--|
| BADW-LRZ | Leibniz-Rechenzentrum der Bayerischen Akademie der Wissenschaften, Germany (3 rd Party to GCS) |
| BILKENT | Bilkent University, Turkey (3 rd Party to UYBHM) |
| BSC | Barcelona Supercomputing Center - Centro Nacional de Supercomputacion, Spain |
| CaSToRC | Computation-based Science and Technology Research Center, Cyprus |
| CCSAS | Computing Centre of the Slovak Academy of Sciences, Slovakia |
| CEA | Commissariat à l'Énergie Atomique et aux Énergies Alternatives, France (3 rd Party to GENCI) |
| CERN | Conseil Européen pour la Recherche Nucléaire |
| CESGA | Fundacion Publica Gallega Centro Tecnológico de Supercomputación de Galicia, Spain, (3 rd Party to BSC) |
| CINECA | CINECA Consorzio Interuniversitario, Italy |
| CINES | Centre Informatique National de l'Enseignement Supérieur, France (3 rd Party to GENCI) |
| CNRS | Centre National de la Recherche Scientifique, France (3 rd Party to GENCI) |
| CSC | CSC Scientific Computing Ltd., Finland |
| CSIC | Spanish Council for Scientific Research (3 rd Party to BSC) |
| CYFRONET | Academic Computing Centre CYFRONET AGH, Poland (3 rd party to PNSC) |
| EPCC | EPCC at The University of Edinburgh, UK |
| ETHZurich (CSCS) | Eidgenössische Technische Hochschule Zürich – CSCS, Switzerland |
| FIS | FACULTY OF INFORMATION STUDIES, Slovenia (3 rd Party to ULFME) |
| GCS | Gauss Centre for Supercomputing e.V., Germany |
| GENCI | Grand Equipement National de Calcul Intensif, France |
| GRNET | Greek Research and Technology Network, Greece |

| | |
|----------------|---|
| INRIA | Institut National de Recherche en Informatique et Automatique, France (3 rd Party to GENCI) |
| IST | Instituto Superior Técnico, Portugal (3 rd Party to UC-LCA) |
| IT4Innovations | IT4Innovations National supercomputing centre at VŠB-Technical University of Ostrava, Czech Republic |
| IUCC | INTER UNIVERSITY COMPUTATION CENTRE, Israel |
| JUELICH | Forschungszentrum Juelich GmbH, Germany |
| KIFÜ (NIIFI) | Governmental Information Technology Development Agency, Hungary |
| KTH | Royal Institute of Technology, Sweden (3 rd Party to SNIC) |
| LiU | Linköping University, Sweden (3 rd Party to SNIC) |
| NCSA | NATIONAL CENTRE FOR SUPERCOMPUTING APPLICATIONS, Bulgaria |
| NREN | National Research and Education Network |
| NTNU | The Norwegian University of Science and Technology, Norway (3 rd Party to SIGMA) |
| NUI-Galway | National University of Ireland Galway, Ireland |
| PRACE | Partnership for Advanced Computing in Europe aisbl, Belgium |
| PSNC | Poznan Supercomputing and Networking Center, Poland |
| RISCSW | RISC Software GmbH |
| RZG | Max Planck Gesellschaft zur Förderung der Wissenschaften e.V., Germany (3 rd Party to GCS) |
| SIGMA2 | UNINETT Sigma2 AS, Norway |
| SNIC | Swedish National Infrastructure for Computing (within the Swedish Science Council), Sweden |
| STFC | Science and Technology Facilities Council, UK (3 rd Party to EPSRC) |
| SURFsara | Dutch national high-performance computing and e-Science support center, part of the SURF cooperative, Netherlands |
| UC-LCA | Universidade de Coimbra, Laboratório de Computação Avançada, Portugal |
| UCPH | Københavns Universitet, Denmark |
| UHEM | Istanbul Technical University, Ayazaga Campus, Turkey |
| UiO | University of Oslo, Norway (3 rd Party to SIGMA) |
| ULFME | UNIVERZA V LJUBLJANI, Slovenia |
| UmU | Umea University, Sweden (3 rd Party to SNIC) |
| UnivEvora | Universidade de Évora, Portugal (3 rd Party to UC-LCA) |
| UPC | Universitat Politècnica de Catalunya, Spain (3 rd Party to BSC) |
| UPM/CeSViMa | Madrid Supercomputing and Visualization Center, Spain (3 rd Party to BSC) |
| USTUTT-HLRS | Universitaet Stuttgart – HLRS, Germany (3 rd Party to GCS) |
| UVSQ | Université de Versailles Saint-Quentin-en-Yvelines, France |
| WCNS | Politechnika Wroclawska, Poland (3 rd Party to PNSC) |

Executive Summary

This deliverable presents the existing and new activities between PRACE and complementary e-Infrastructures with the aim to build an integrated Europe-wide e-Infrastructure for our users [1]. The focus has been on the continuation and further development of the collaboration on security and data management topics.

PRACE has been collaborating on security with other e-Infrastructures from its beginning. This includes cooperation with major European organisations (EGI [2], EUDAT [3], GÉANT [4], Human Brain Project [5]), as well as with Centres of Excellence (CoEs [6]) such as EoCoE [7], POP [8] or BioExcel [9] and CoeGSS [10]. During the first call for projects, the PRACE council reserved 0.5% of the total calculation resources available for the use of the initial nine CoEs. The same has applied to the second call, which began in January 2019, with two more new centres: HiDALGO [11] and EXCELLERAT [12]. There are also some ongoing preliminary discussions with worldwide consortiums (XSEDE [13] and RIST [14]).

Workshops attended by representatives from EGI, EUDAT, GÉANT, NRENs, PRACE as well as US participants, are held regularly to organise these partnerships. The follow-up activities of this forceful collaboration are organised in working groups, which cover specific topics such as networks, authentication and security, data storage, training, etc. that are coordinated by an informal Steering Committee, in which PRACE is actively engaged. PRACE has a long history of collaboration with these e-Infrastructures and uses the services of these e-Infrastructures in various areas of interest to it.

PRACE collaborates with EUDAT for data services, which offers a service catalogue [15] complementing PRACE in order to fulfil user needs during the full research data life cycle. During the past years, some integration in the areas of data transfer, storage and sharing has been implemented as data pilots were run as part of the DECI-13 call and the DECI-14 call. A Data Management Plan template [16] has been used to ease requirements gathering during the early phase of each pilot and its systematic use is now highly recommended. The need for cross PRACE and EUDAT trainings has been identified. This collaboration with EUDAT is increasing, with delivering some sessions at PRACE training events and vice versa.

Regarding authentication and authorisation services, PRACE is considering to use services provided by the Identity Provider (IdP) federations which exist for educational organisations in several countries and which are federated in eduGAIN [17]. Working with external IdPs can enhance the security of the PRACE infrastructure because users will not have to use separate credentials for PRACE access. This task follows other activities in this area, especially those of the AARC [18] project which is coordinated by GÉANT.

Within the framework of the European Cloud Initiative, a desire has emerged to federate world-class supercomputer capabilities with high-speed connectivity and leading-edge data and software services for science, industry and the public sector. This is how the European Data Infrastructure (EDI) project was born, where GEANT and PRACE have a fundamental role to play.

For security alert and incident management, PRACE has been collaborating with EGI and EUDAT for a long time. This cooperation will continue through the WISE [19] community initiative.

1 Introduction

One of the crucial directions given as part of the H2020 Programme (EOSC/EDI [20]) is to design a European Area of Research, Science and Technology to foster scientific excellence [21]. The provision of world-class research infrastructure, as well as priority and seamless access for researchers to this type of infrastructure, are therefore essential in achieving this objective. In this integrated vision, a scientific workflow should be able to use services from PRACE as well as from other European providers such as EUDAT, EGI or GÉANT, or even from other continents (XSEDE and RIST for example).

This report intends to provide an overview of the collaborative activities underway at the end of the PRACE-5IP project and to provide a state-of-the-art on potential and future collaborations. The overall objective is to implement collaboration as much as possible on the basis of user-driven pilot projects.

In Section 2, we present collaborations related to scientific networks for intensive computing. We provide an overview of PRACE's collaborations, by partnership, in the operational area and the key points of the collaborations are further developed in the subsections. There is also a focus on vertical collaboration, the interoperability with EUDAT data services, starting with a presentation of the benefits provided to PRACE users by using EUDAT data services. We then describe how we identified the relevant pilots and finally, we summarise the collaboration we have established around HPC and data management training.

A status about ongoing and future activities with Centres of Excellence is provided in Section 3. We explain why these collaborations are important, how these CoEs contribute to the success of the EU HPC ecosystem, what resources are being used and deployed, and the current state of thinking towards a fruitful development of these collaborations.

Section 4 describe security-related activities which play a key role in the sustainability of collaborations. We explain our involvement in the WISE community and provide an update on the ARCC2 EU project and on the Trusted Introducers services. Section 5 closes the report with final conclusion including a summary and vision of future co-operation.

2 Collaboration with other e-Infrastructures

This section presents existing and prospective collaborations with various PRACE partners. Formal agreements such as a Memorandum of Understanding (MoU) and areas of collaboration are then described by partner but also by project when it involves several stakeholders.

2.1 GÉANT

The collaboration with GÉANT has a long history since the network dedicated to PRACE, which began in the DEISA project in 2006, was entirely based on GÉANT's footprint. PRACE took over from the network implemented. A close cooperation has been in place since then, which has been detailed in the previous PRACE-4IP reports [22] [23].

Recently, a meeting involving PRACE-GÉANT representatives took place in Bologna (February 2019), where further collaboration opportunities were reviewed such as PRACE network, identification of areas of synergy, authentication and authorisation infrastructure with AARC project and eduGAIN, and training (workshops) were among the topics discussed.

As part of the prototyping activities, PerfSonar's experimentation could start during PRACE-5IP. The perfSONAR Toolkit [24] is a widely-deployed test and measurement infrastructure a) designed to facilitate single shot installation and configuration of network measurement tools and b) that support deployments of all sizes. This infrastructure features a set of tools that gather and share metrics of interest about a target network path including e.g. achievable bandwidth, jitter, and latency. The task of gathering network measurements relies on methods that sample the medium and requires stable and predictable end-to-end communication patterns to complete the operation.

PRACE and GÉANT agreed on installing this software package in the PRACE MDVPN network to measure throughput and latency between Tier-0 and Tier-1 HPC systems.

As a first test, it was agreed to install the perfSONAR testpoint software (clients) on systems at PRACE sites subnets, which are also participating into the HBP HPC infrastructure, i.e. BSC, CINECA, ETH Zürich/CSCS and JUELICH. More PRACE partners willing to participate in a first test pilot have been welcomed to join. For the time being testpoint software has been installed at IT4Innovations and KIFÜ additionally. The perfSONAR test environment has been setup including a MaDDash server [25] providing a configuration file, defining, which tests should be scheduled, an ESMOND database [26] storing the test results and an Apache frontend (monitoring portal), which allows authorised users to see the results of the tests completed.

At time of this writing, CINECA, JUELICH, and KIFÜ have been already able to read the configuration file and schedule tests accordingly.

It is currently investigated why the results produced by KIFÜ cannot be stored into the central database. A second server in JUELICH can read the configuration file, but does not schedule tests as requested. Both problems are investigated by GÉANT software developers currently.

ETH Zürich/CSCS and BSC are in the process of setting up their servers. IT4Innovations has to wait until their system can be updated to a new OS software version. The current OS version

(CentOS 6) is not further supported by GÉANT and it is recommended to upgrade to CentOS 7 before installing the software.

Currently, the monitoring portal does not have any access restrictions. In the future, when in production, it is planned to setup access rights, as soon as a PRACE IdP and Attribute server have been implemented. Figure 1 shows the current test scenario for the PRACE perfSONAR installation.

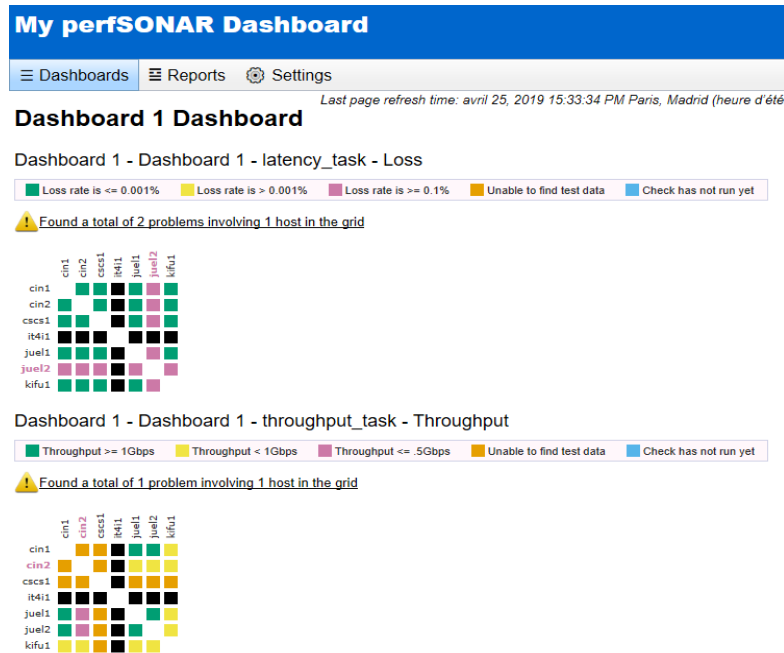


Figure 1: PerfSONAR Dashboard [27]

2.2 EUDAT

It is well-known that the capacity to couple data and compute resources together is a key aspect to accelerate scientific innovation and advance research. This is why, since the beginning of the collaboration between PRACE and EUDAT, two MoUs have been signed between these two infrastructures. The first agreement, signed in March 2014, allowing EUDAT centres to use the PRACE network to transfer data within the infrastructure for replication purposes or to improve accessibility. With the advent of the Horizon 2020 projects, the objective has been revised to offer seamless integration of data services between PRACE and EUDAT and the second MoU has been signed for this purpose in mid-2016. In addition to the network usage, this MoU covers the definition of governance and support processes, the identification of training and methods, and the implementation of pilot projects.

2.2.1 Leveraging EUDAT services

One of the objectives of PRACE-5IP WP6 is to increase PRACE's service portfolio and to provide PRACE users with additional data management services. EUDAT is seen as a perfect fit as it

provides a collection of Data Management services [15], while PRACE is focused on HPC resources.

In this context, WP6 implemented selected pilot projects within the framework of the PRACE/DECI. These projects took advantage, on the one hand, of the computational performance offered by PRACE in the data production phase and on the other hand the long-term storage and preservation of data offered by EUDAT in the post-production phase. The most widely used EUDAT service is B2STAGE, due to its ease of use and reduced preparation times. B2SHARE is also used to publish information material, publications, reports on the results of PRACE resource allocation, white papers, guides to good practice, etc.

The close integration of PRACE and EUDAT services are described on Figure 2.

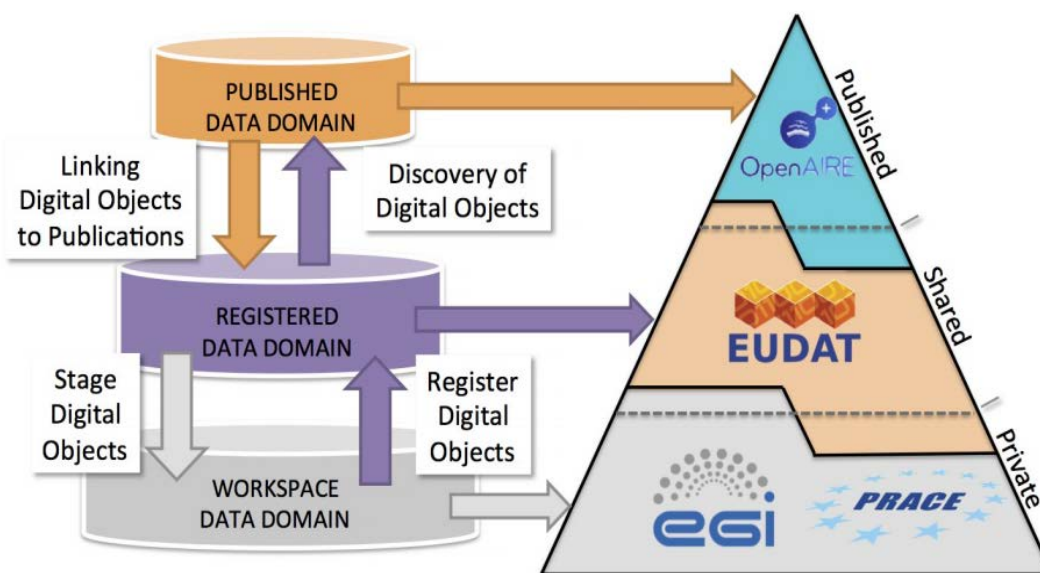


Figure 2: EUDAT Data Domain modelled on the ANDS [28] Data Curation Continuum

Data Management Plan (DMP) template elaborated by EUDAT [16] has been used by such pilot to gather the user requirements in order to understand the expected use cases and the nature of the data, and ensure the full data life cycle is taken into account. To meet the data requirements of these pilot projects, we have identified some existing EUDAT services such as data transfer, publication and search of data and metadata, or data preservation. It also allowed us to describe new services such as storage in a workspace or data processing on a storage site.

It is crucial to work towards a standard service approach for such outsourcing. This will ensure an easy uptake and clear agreements for users and communities as well as inclusion into scientific workflow. Seamless services integration between e-Infrastructures must rely on robust technical solutions.

2.2.2 DECI-13 and DECI-14 Pilots

The DECI-13 and DECI-14 calls are issued as a pilot calls jointly with EUDAT. They offer an opportunity to allocate compute and data resources to projects that also have needs for storage capacity and associated data services. Applicants may apply for both the DECI and EUDAT calls at the same time or may apply to just one of the calls.

Five projects had been selected as pilots during the DECI-13 call, see Table 1:

| Project Acronym | Project Title | Field | Data req. in TB | PI country | Exec site | EUDAT site |
|-----------------|---|-------------------|-----------------|------------|----------------|------------|
| HybTurb3D | Hybrid 3D simulations of turbulence and kinetic instabilities at ion scales in the expanding solar wind | Astro Sciences | 140 TB | IT | SURFsara | CINECA |
| MULTINANO | Multiscale simulations of nanoparticle suspensions | Engineering | 30 TB | IT | RZG | CINECA |
| HiResClimate | High Resolution EC-Earth Simulations | Earth Sciences | 150TB | IE | KTH | EPCC |
| AFiD | Effect of rotation and surface roughness on heat transport in turbulent flow | Engineering | 11TB | NL | EPCC | SURFsara |
| CHARTERED | Charge transfer dynamics by time dependent density functional theory | Materials Science | 30TB | SE | IT4Innovations | KTH/PDC |

Table 1: DECI-13 call data pilots

The use cases of these pilots were conducted from April 2016 to March 2017 and are described in the PRACE-4IP D6.6 deliverable. Initial feedback has been presented during the final EUDAT conference in Porto in January 2018, and the final evaluation has been reported in the EUDAT deliverable published in March 2018 [29].

Three out of the applications requested for access to EUDAT resources during the DECI-14 call within the PRACE Optional Programme. At the end of the DECI appraisal process, only two candidates have been selected. One of them is the continuation of data pilot that ran as part of DECI-13. Both started in April 2017.

| Project Acronym | Project Title | Field | Data req. in TB | | PI country | Exec site | EUDAT site |
|-----------------|--|-------------------|-----------------|--|------------|----------------|---------------------|
| subGridEoR | The effects of small-scale structure and halo stochasticity on Cosmic Reionization | Astrophysics | 100TB | | UK | SURFsara | Commitment withdraw |
| CHARTERED2 | Charge transfer dynamics by time dependent density functional theory | Materials Science | 30TB | | SE | IT4Innovations | KTH/PDC |

Table 2: DECI-14 call data pilots

2.2.2.1 CHARTERED2 Data Pilot

2.2.2.1.1 Project and actors

The CHARTERED2 project had been awarded as part of the DECI-14 call. This project is in the field of Materials Science, focusing on “Charge transfer dynamics by time dependent density functional theory”, and had initially been selected as part of the DECI-13 call. The project team is based in Uppsala University (Sweden) and the PRACE execution site was IT4Innovations (Czech Republic) and the EUDAT site was PDC/KTH (Sweden).

The data management plan had been initiated during the first phase of the pilot, and completed iteratively along some regular calls. The data sharing section of the document is not yet filled, as the results of the simulations need to be evaluated before publication. This will be done at a later stage.

The data flow diagram in Figure 3 describes the objective of the pilot based on the requirements gathered.

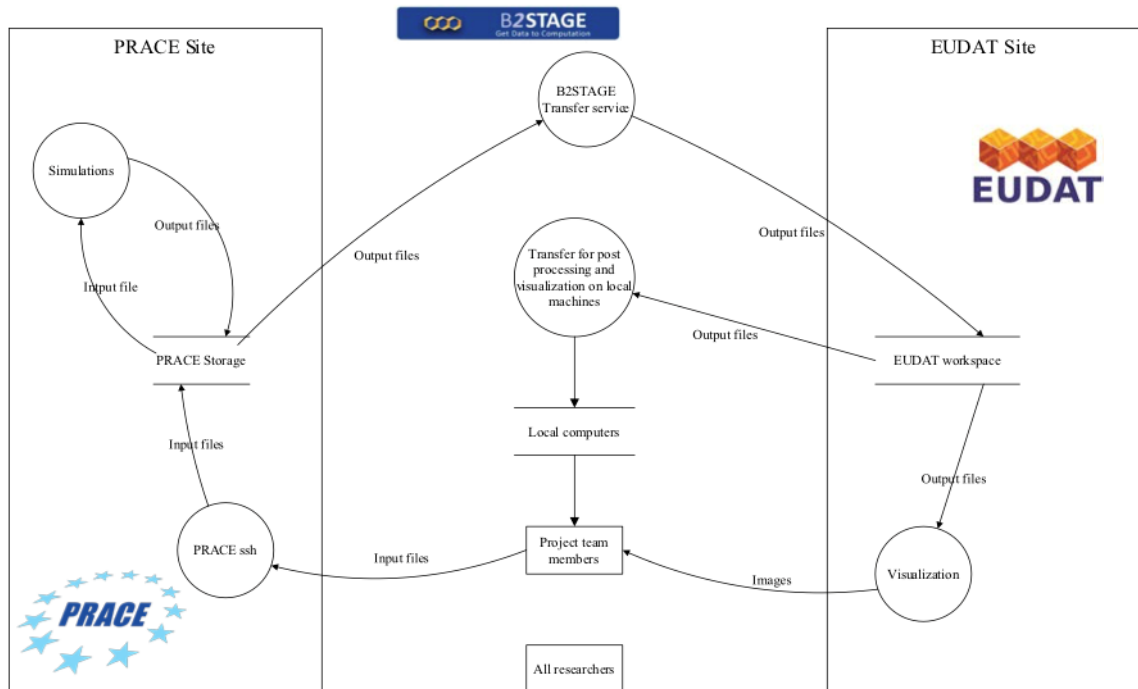


Figure 3: Data flow diagram for CHARTERED2 data pilot

The data requirement for EUDAT pilot was 20 TB during and after the PRACE project. As requested, the EUDAT site was PDC/KTH (Sweden).

To start with, there were two logins per person: the PRACE login and the PDC/KTH login. All five team members already had PRACE accounts associated with the project and PDC/KTH accounts.

As PDC/KTH is not connected to the PRACE network, the public network between PDC/KTH and IT4Innovations was used. The latter is connected to public Internet via redundant 40 GB/s connection on top of which there is also an IPsec 1 GB/s tunnel provided to securely connect to PRACE network. Over both connections, standard transfer tools like SSH/SCP, GSI-SSH/GSI-SCP and gridFTP are supported. All these transfer tools provide encrypted file transfer and thus securing the contents transferred between the IT4Innovations and KTH/PDC sites. B2STAGE was installed in PDC including up-to-date GridFTP and relevant certificates. For data transfer, GridFTP is used because it is the solution that works behind EUDAT B2STAGE. Between EUDAT and local computers, GridFTP is used since a part of the post processing is performed at user's location.

2.2.2.1.2 Current status

IT4Innovations prolonged CHARTERED2 project run quota until 31 August 2018, so the project runs have now ended and data should be transferred from IT4Innovations back to Sweden. Part of the data pilot was to develop the mechanism how to transfer data continuously from IT4Innovations and it was never expected to keep it safe and for a longer period of time at IT4Innovations. CHARTERED2 data pilot has now over 20 TB of data at IT4Innovations. Unfortunately, during

CHARTERED2 data pilot runs at IT4Innovations, EUDAT2020 project ended and PDC/KTH no longer operates B2STAGE (their EUDAT storage services were decommissioned). National storage resource in Sweden is now Swestore (operated by SNIC, The Swedish National Infrastructure for Computing). CHARTERED2 project team has applied for the storage allocation in Swestore. However they have not yet received the storage allocation decision from the Swestore and CHARTERED2 project data is still at IT4Innovations.

2.2.2.2 SubGridEoR Data Pilot

As a data pilot, we must consider this project failed. The project was proposed and granted during the DECI-14 call. In its proposal, it had indicated that the project wanted to make use of the option given to make use EUDAT storage resources to preserve and share data produced by the computations. More specifically, the proposal indicated that it needed a volume of about 100TB on B2SAFE storage for two years. In the granted proposal the primary investigator had also mentioned two preferred EUDAT partner sites.

The computational part of the project started end of April 2017 on Cartesius at SURFsara. While the project was under way, it became clear that within EUDAT, the site selection process to actually provide the required volume of B2SAFE resources was not yet finalised. The communication and decision making, on which EUDAT partner would provide the storage resources, was rather difficult from the outset. The situation was escalated to the EUDAT Work Package leader, to bring about a solution. In February 2018, it was understood that STFC Hartree was available as the EUDAT site to offer space to accommodate the data and would expose a B2SAFE end point for data ingestion using the GridFTP.

However, in July 2018 STFC Hartree communicated that the B2SAFE service that they had established was at a much smaller scale and was already being used or reserved for other projects. Consequently, it would not be able to provide the storage capacity needed for subGridEoR project in the coming months.

On Cartesius, the project was already extended, not because of the storage uncertainties, but due to the delay caused by some computational problems which occurred at the start. The extended project definitely ended on 1 September 2018. Some of its data have been secured elsewhere, at the home site of the primary investigator, but within the timeframe of the pilot project, no data have been transferred to an EUDAT repository.

2.2.2.3 Analysis of the outcomes

Although these two pilot projects did not succeed in following their data management plan to the end, they allowed us to highlight some problems in using EUDAT's storage services. A decommission of the B2STAGE service at the initial EUDAT site prevented the transfer to the initial storage location for the CHARTERED2 project. Work on service continuity must be considered. And problems in terms of the required storage volume have clearly shown that the process of selecting the EUDAT site to effectively provide the required volume of B2SAFE resources needs to be improved.

2.2.3 Training activities

While working on the data pilots described above, it appeared that PRACE users are usually small groups (a project and not a community) who produce large volume of data. In this context and in the light of the recent evolutions in the domain of data management, it is necessary or very useful for some users to develop these skills and to be trained to good practices. A working group involving PRACE and EUDAT training managers, meeting via video conference, was established to facilitate this.

Several training materials, such as webinars, presentations and briefing papers [30] have been created and are available on the EUDAT website. The objective is to present the challenges of research data management, the services required throughout the data life cycle, as well as the European Commission's recommendations on this subject.

To strengthen the collaboration, the workgroup has agreed on the cross reference of training websites from both e-Infrastructures. Training events from the partner are also advertised on each other's mailing list. Finally, it was agreed to include some PRACE content into EUDAT trainings and vice versa. The calendar for this collaboration can be seen below:

- During PRACE 2017 Spring School joint event with VI-SEEM, Cyprus, EUDAT led a session about 'Data Management Plans – EUDAT best practices and case study' as part of the Data/Computational Services for Scientific Communities stream (25 to 27 April 2017).
- During the HPC Summit Week 2017, Barcelona, PRACE contributed a workshop about "Coupling HPC and Data Resources and services together", as part of a joint event during the PRACEdays17 (18 May 2017).
- During ISC 2017 conference, Frankfurt, PRACE organised the 3rd session in the series of the education and training related workshops (22 June 2017). A EUDAT training team member participated to the event.
- During the EUDAT & Research Data Management Summer School, Heraklion, partners involved in PRACE delivered several sessions about HPC data management (July 2017).
- During the PRACE joint event with BioExcel, Stockholm, focusing on ARM & EUDAT performance optimisation and data management tools, EUDAT led a session on DMP good practices (May 2018).
- As part of the training program of the H2020 PRACE project, training on the Data Management Plan was given at CINES in Montpellier (February 2019).

Now that this collaboration has started, we have and will use feedback from the data pilots and the trainings to refine the training content.

2.3 EGI

EGI has been involved for years with PRACE in different workgroups related to operational security and trust models. In order to further develop this collaboration, an MoU has been signed in March 2016. In addition to security related topics, it targets the ability to build an integrated service catalogue and more interoperability to support data driven science. At this stage, security-related activities are still evolving (see section 4), as well as activities on interoperability and the

integrated service catalogue (particularly via e-InfraCentral and eduGAIN as explained in the following section 2.4).

2.4 e-InfraCentral

eInfraCentral is a CSA Project [31], aiming to create the conditions to ensure that, by 2020, a broader and more varied set of users (including industry) discovers and accesses the existing and developing e-Infrastructures services in Europe.

The eInfraCentral project is a consortium of nine entities (public and private), including five flagship European e-Infrastructures (EGI, EUDAT, GÉANT, OPENAIRE, PRACE). To fulfil its mission, eInfraCentral is building a common catalogue of European data, compute and network services to be included in a portal being developed [32]. In more general terms, eInfraCentral aims at becoming an important EOSC gateway, an aggregator of services through an access portal to pan-European Open Science resources, supporting multiple marketplaces to promote their services to end users. To this end, there is a collaboration with the EOSC-hub project [33] which has common goals with the eInfraCentral project.

Besides that purpose, and because it is perceived that there is a clear need for all stakeholders to understand how to assess the use, operation, innovation and impact of e-Infrastructures in Europe, the portal developed by eInfraCentral also includes the monitoring of KPIs that focus on availability, quality of services and user satisfaction.

In order to have seamless integration of the Service Catalogue of each e-Infrastructure in the portal, eInfraCentral seeks to achieve standardized service descriptions, aggregation of service data and common presentation, as well as standard methods for exchanging service-related information among those service providers. This is implemented by a specially developed API which allows for this exchange of information between the portal and the websites of the different e-infrastructures, the ultimate goal being the possibility of mining the relevant data from those sites and updating the relevant information automatically.

PRACE has been involved in this effort, both in the managerial and policy side, as well as in the more technical aspects. Regarding the former, PRACE has contributed to define several types of services to be included in the overall catalogue, the categories of KPIs to be used, including their descriptions, and several other similar tasks, in particular participating in discussions about the EOSC portal [33], which would take advantage of the work done within eInfraCentral. The technical tasks are done by KIFÜ (Hungary), the PRACE partner in charge of the PRACE website. KIFÜ has been testing the aforementioned API regarding the exchange of PRACE services content between the PRACE site and the eInfraCentral portal and will do the same regarding KPIs information.

2.5 Human Brain Project

The PRACE-HBP MoU was signed in November 2014 and focuses on the PRACE network usage by the HBP community. The access to the PRACE infrastructure seems obvious as the main HBP HPC partners (BSC, CINECA, CSCS, and JUELICH) are already members of PRACE. Since April

2018, TGCC (Très Grand Centre de calcul) of CEA has joined the Human Brain Project's High Performance Analytics and Computing (HPAC) Platform. With this new partner, all PRACE Hosting Members are now in the Human Brain Project.

2.6 XSEDE / RIST

Research teams are increasingly distributed on a global scale. To meet their computational requirements, PRACE, XSEDE and RIST, three of the main players in the global high-performance computing ecosystem, are working to develop some collaboration. To do this, an MoU was initiated by senior management from the three advanced research computing institutions following a meeting during SC'16 in Austin, with the objective to join forces to increase information exchange and research collaboration.

2.7 Collaboration with CERN

During the PRACE-5IP project, WP6 started to contact CERN representatives in order to establish a collaboration based on the identification of mutual needs in terms of computation and services in the HPC ecosystem.

In collaboration with T6.2, we had a preliminary meeting in CERN premises in Geneva in September 2017. Here, with the presence of the PRACE-5IP PMO and the main leaders of the CERN IT services, we presented the PRACE roles in the HPC European ecosystem. We discussed about way to promote excellence in the computational research domains and many activities in the field of training and enabling of science. On the other side, CERN representatives conveyed the computational needs related to the experiments at present or in the near future in the laboratories (with a special remark to the needs that will be related to the High-Luminosity LHC, codenamed HL-LHC, experiment). At the end of this meeting, PRACE and CERN decided to have a follow-up meeting involving more people from the CERN experiments.

This follow up took place again in Geneva, at the CERN labs, in October 2018 in the form of a workshop with the participation of the scientists involved in the experiments and on the data analysis. On the other side, there were the PRACE-5IP PMO together with the PRACE Managing Director and the PRACE5IP WP6 leader.

The first part of this workshop focused on the PRACE mission and structure, presenting the opportunities of training and joint PRACE-CERN activities. We also focused on the CERN and SPS experiments, sharing the CERN-IT plans for the future as well as specific requirements for Machine Learning workflows of interest for the LHC experiments.

This meeting was productive, showing strong willingness to collaborate on all sides. The meeting had an ambitious list of goals, most of which were met:

- Present how PRACE works and how we could establish a connection with CERN;
- Understand the mechanisms of the PRACE resource allocation, scheduling, I/O and data serving, included the services for AAI;

- Discuss the future of PRACE in the context of the other EU initiatives in the HPC ecosystem, including the plans for future investments in the hardware facilities;
- Understand what can be the exploitation of HPC related to the HEP in terms of hardware and software together with PRACE;
- Present the current usage of the HPC resources related to the HEP experiments, most of them are now based on individual contacts between CERN and European computing centres;
- Discuss potential and/or possible changes in the infrastructure and application side to facilitate the exploitation of the HPC infrastructures within the PRACE model.

During the discussion, it emerged that there are significant differences in scheduling, access and other services between the PRACE HPC community and the CERN LHC usage model. It is not realistic that only one of the two communities could fill the existing gap. It is worth recognizing that, improving the scientific applications to be better tuned for HPC environments and pushing the HPC centres to adopt more flexible exploitation models, would bring the communities much closer.

A few interesting challenges were identified:

- LHC and other data intensive experiments, such as SKA, are often open-ended multi-year projects, which need predictable computing resources that cannot be easily accommodated by the PRACE annual proposal-driven allocations;
- Most of the LHC/HEP applications are typically “embarrassingly parallel”, which means that the interconnect of the HPC infrastructure is typically not used at all;
- A common set of interfaces/procedures for AAI, resource allocation and data management should be adopted by the HPC centres to reduce the burden related to differences in such practices;

During the wrap-up of the workshop, some strategical and technical actions were proposed in order to overcome the challenges discussed above. Some of the proposed activities could be considered relevant also for the AEGIS project and the SKA experiment.

Proposed activities for future collaboration with CERN are reported in Table 3. We are planning to continue and consolidate the link between PRACE and CERN during the PRACE 6IP project.

| Strategic | Technical |
|--|--|
| 1. Develop a three-way agreement between CERN, SKA and PRACE to explore long-term cooperation to support the LHC and the SKA science programmes. | 1. Definition and execution of a pilot project for an HTCondor overlay for HPC resources in PRACE based on similar work at CERN. HPC resource allocation could then be performed generically for the LHC community via an HTCondor pool. |
| 2. Survey of current LHC experiment activities using PRACE resources. Document | 2. Definition and execution of a data federation demonstrator: the goal is to |

| | |
|--|---|
| successes and challenges. Develop a list of priorities for a development project to make efficient use of PRACE sites and resources for as many LHC workflows as possible. | demonstrate data delivery at run-time through the HPC firewalls at an incoming rate sufficient to efficiently operate at scale HPC processing resources for data intensive workflows |
| 3. Development of services and tools for LHC and SKA workflows. Assessment of e-Infrastructures like FENIX (Fusion Engineering International Experiments) | 3. Definition of a programme to demonstrate that local storage at HPC sites can record data at production scale from high-output workflows (reconstruction and simulation). This demonstrator should also verify that export from the HPC facilities to remote custodial storage can be sustained at a level which permits continuous operations. |
| 4. Develop a software-driven initiative to tackle the LHC and SKA needs in terms of software optimisation and performance on HPC resources. | |
| 5. Creation of training programmes tailored to the LHC experiments' adoption of HPC architectures and performance optimisation of high-end accelerators such as GPU and FPGAs. | |

Table 3: Proposed activities for future collaboration with CERN

2.8 Collaboration to support Globus based services

Data transfer services within PRACE have relied on the Globus Toolkit, an implementation of the GridFTP extension of the file transfer protocol (FTP). The Globus Alliance, which was in charge of the development, announced that support for the toolkit would end after 31 December 2017 and that important security patches would be provided until the end of 2018. In order to ensure further availability of the discontinued toolkit, the Grid Community Forum (GridCF) [34] was founded. It is a global community that provides support for the Grid Community Toolkit (GCT), which is an open source fork of the discontinued toolkit. The main goal is to maintain the quality and security of the Grid Security Infrastructure (GSI) and the GCT, to provide a framework for external interested parties that wish to contribute to software development and to support collaboration with other grid organisations. The GridCF is well organised, having the project management committee and adopted standards and best practices.

First GCT precompiled packages were available for testing purposes in September 2018 followed by the stable ones in October 2018. The current stable version is available either as source or via precompiled packages for Debian and Ubuntu, for RHEL and RHEL compatible Linux distributions (from the EPEL repo), for Fedora and for SLES and OpenSUSE (from the OpenSUSE

Build Service). Source code is managed using git version control system. GCT contains three major components:

- Grid Security Infrastructure, which is based upon the concept of X.509 proxy certificates;
- GridFTP, or an implementation of the GridFTP protocol;
- MyProxy, which is a credential management service tailored for GSI.

Before support for important security patches for Globus toolkit had ended, PRACE sites have been advised to upgrade to the GCT to benefit from the included community patches and to switch to a maintained drop-in replacement for the Globus Toolkit in time. As of 14 March 2019, 8 partners have upgraded their data transfer nodes to GCT. Table 4 is also available at the PRACE Wiki pages under the request for comments #282 [35].

| Site | Status | Comment |
|---------------|----------------|--|
| BSC (es) | | |
| CASTORC (cy) | Not applicable | System decommissioned |
| CEA (fr) | | |
| CINECA (it) | done | EPEL |
| CINES (fr) | | |
| CSC (fi) | | |
| CSCS (ch) | | |
| CYFRONET (pl) | | |
| EPCC (uk) | | |
| FZJ (de) | done | epel |
| GRNET (gr) | Done | EPEL |
| HLRS (de) | done | |
| ICHEC (ie) | Done | 2019/01/14. EPEL. |
| IDRIS (fr) | | |
| IUCC (il) | | |
| LRZ (de) | Done | 18/12/2018 |
| NCSA (bg) | | |
| NIIF (hu) | | |
| PDC (se) | | |
| PSNC (pl) | done | changed from Globus-Toolkit-6-Stable to epel |
| RZG (de) | | |
| SURFSARA (nl) | done | 3/12/2018, using epel repository |
| UHEM (tr) | | |
| UIO (no) | | |
| UL (si) | | |
| VSU-TUO (cz) | done | 10/01/2019, EPEL, Anselm & Salomon |
| WCSS (pl) | | |

Table 4: Status of upgrade to GCT among PRACE partners

3 Collaboration with CoEs

Nine Centres of Excellence (CoEs) for computing applications have been funded as part of the EC Horizon 2020 programme. The WP6 team had several meetings with their representatives (as reported in previous reports, e.g. PRACE-4IP D6.5 [22] [23]) in order to understand their needs, give them access to the relevant PRACE services and organise joint events. The list of CoEs is shown Table 5.

| Coordinator / Organisation | Country | Acronym | Title |
|---|---------|------------|---|
| COMMISSARIAT A L'ENERGIE ATOMIQUE ET AUX ENERGIES ALTERNATIVES | FR | EoCoE | Energy oriented Centre of Excellence for computer applications |
| KUNGLIGA TEKNISKA HOEGSKOLAN | SE | BioExcel | Centre of Excellence for Biomolecular Research |
| MAX PLANCK GESELLSCHAFT ZUR FOERDERUNG DER WISSENSCHAFTEN E.V. | DE | NoMaD | The Novel Materials Discovery Laboratory |
| CONSIGLIO NAZIONALE DELLE RICERCHE | IT | MaX | Materials design at the eXascale |
| DEUTSCHES KLIMARECHENZENTRUM GMBH | DE | ESiWACE | Excellence in Simulation of Weather and Climate in Europe |
| UNIVERSITY COLLEGE DUBLIN, NATIONAL UNIVERSITY OF IRELAND, DUBLIN | IE | E-CAM | An e-infrastructure for software, training and consultancy in simulation and modelling |
| BARCELONA SUPERCOMPUTING CENTER – CENTRO NACIONAL DE SUPERCOMPUTACION | ES | POP | Performance Optimisation and Productivity |
| UNIVERSITAET POTSDAM | DE | COEGSS | Centre of Excellence for Global Systems Science |
| University College London (UCL) | UK | CompBioMed | Centre of Excellence for Computational Biomedicine |

Table 5: List of Centres of Excellence

The importance of establishing a link with the Centers of Excellence became more relevant than in the past, due to the importance of their (CoEs) role in EuroHPC era. In the past, Work Package 6 started to establish connections with some CoEs, in particular with EoCoE and PoP. Work Package 6 had several meetings with the aim to establish MoUs in order to share common services or to fulfil reciprocal operational needs. During some months of PRACE 5IP, part of this activity, was put on hold in the wait of the results of a new H2020 call for the renewal of the CoEs. Then, in the second part of 2018, with the new CoEs this work was resumed and, in particular, to foster this collaboration between PRACE and CoEs, a meeting, co-organized with EXDCI, took place in October 2018 in Brühl (DE).

3.1 PRACE-EXDCI-CoEs meeting in Brühl, Germany

In order to give new strength to the collaborations between PRACE, CoEs and FET projects, in October 2018, PRACE and EXDCI organised a two-day workshop in Brühl. The aim of this workshop was to start or revamp a spirit of collaboration among the different EU funded initiatives, identifying common needs and areas of collaboration.

During this meeting, WP6 coordinated the organisation of two sessions for the “Operational Services” in which several representatives from different CoEs and FET took part. In the first of the two sessions, we had short presentations of the activities performed by each of the represented projects: PRACE (F. Affinito), DECI (C. Johnson), HiDALGO (M. Gienger), CompBioMed (G. Pringle), MaX (F. Affinito), ESIWACE (K. Serradell), and ECOSCALE (Y. Papaefstathiou). The idea of these small talks was to demonstrate the work performed (or to be performed) under each project to the representatives of the other projects. And, also to present their requirements in terms of computational resources and/or services. In the second session, the following day, we dedicated the whole time to an open discussion to individuate the needs of each project and the areas in which the work could be shared. A significant part of the discussion was dedicated to understand the limiting factors to the collaboration among the different projects, including PRACE.

Many interesting points emerged during these two sessions, in particular:

- CoEs need computational resources. PRACE is offering a part of the allocated budget to the CoEs, but this budget is sometimes considered too little or too difficult to be obtained
- This need of computational resources sometimes does not fit the allocation model of PRACE: the “opportunistic model” offered by commercial cloud services seems more suitable in some cases.
- In any case, there is no silver bullet: the requirements change for different CoEs. In some cases, the cloud services model is a better fit. In some other, the resources offered by PRACE could not be sufficient.
- Use cases coming from the CoEs could offer the opportunity to deploy new services. A collaboration in this sense could give back relevant and useful feedback.
- For some of the CoEs, in particular ESIWACE and CompBioMed, the concept of “Urgent Computing” seems interesting. In any case, it should be further clarified in which use cases it is more appropriate to call that “High Availability” than “Urgent Computing”. This is not the only case where the wording in different contexts (i.e. different projects) can assume different meanings.
- In order to tackle the different needs of resources, it emerges that a push towards a federation of the resources and the services is necessary.

Unfortunately, due to the lack of time discussion about the integration of data services could not take place.



Figure 4: PRACE-CoEs-FET HPC-EXDCI Workshop – October 2018

The discussion during the two days was very active and lively. This is a signal that there is a huge potential of interaction among the different projects, in particular the CoEs, which just needs to be addressed.

One important thing that emerged during this discussion, and that concerns very much this task in PRACE-5IP, is the absence (or the difficulty to individuate) of a point of contact for PRACE. This is probably the most important action for PRACE. Immediately after the Brühl meeting, discussions started in WP6 in order to set up a quick and visible channel of communication addressed to facilitate the interaction of CoEs with PRACE.

This meeting triggered many initiatives aimed at establish a consolidated link between PRACE and other EU-funded projects. PRACE set up mailing lists, concerning the different topics discussed in Brühl to keep in contact the representatives of the CoEs and FETs with PRACE key-persons. Furthermore, we started to discuss how to create a single point of contact between PRACE and external projects/initiatives in order to facilitate and speed-up the collaboration.

A follow-up of this meeting is scheduled during the EuroHPC Summit Week in Poznan (May 2019).

3.2 BioExcel Centre of Excellence

The aim of the BioExcel CoE is to enable better science by improving the most popular biomolecular software and spreading best practices and expertise among the communities through

consultancy and training. It acts as a facilitator, a central hub for biomolecular modelling and simulations.

Following the success of the first edition of the Seasonal School (April 2017) [36], the second edition is being prepared for June 2019 at KTH [37], jointly organised by PRACE and BioExcel. The objective is to acquire a comprehensive understanding of the use of the most popular and well-known applications in the field (Gromacs, Amber, NAMD, VMD), to understand their scalability and performance, while avoiding potential issues and to learn best practices regarding their use on HPC systems.

3.3 Energy oriented Centre of Excellence

EoCoE is a hub coordinating a pan-European network of partners, which are strongly engaged in both scientific simulation and energy fields. It aims to foster and accelerate the European transition to a reliable and low carbon energy supply using HPC.

Following the first "Scientific Applications towards Exascale" workshops held at CINES [38], a second one was held in June 2018 at the Maison de la Simulation in the form of a "Performance Engineering Hackathon". This time, the training consisted of lectures on performance engineering methodology, practical sessions for applying the methodology on trainee's intensive computing kernels.

3.4 Performance Optimisation and Productivity Centre of Excellence

The POP CoE in Computing Applications provides performance optimisation and productivity services for academic and industrial code(s) in all domains. The POP CoE, with its service and training activities, has had a wide impact in all areas of research and industry. It has provided researchers and industry with access to HPC expertise that allows them to be more productive and achieve scientific and industrial excellence. It has also improved the competitiveness of the centre's customers, by generating a tangible return on investment (ROI) in terms of savings, elimination of waste, errors and delays, by making their applications leaner and issue-free.

Following the success of the first phase of this project, a second phase - POP2 - continues for an additional period of three years (2019 to 2021). A proposal was submitted and accepted by the partners for the INFRAEDI-02-2018 call: HPC PPP - Centres of Excellence on High Performance Computing. The POP2 project team now has two additional expert groups:

- The Performance Tools (PT) team at UVSQ (France) [39], which has been working since 2004 on performance optimization tools and methodologies, including its Modular Assembly Quality Analyzer and Optimizer (MAQAO) suite of performance analysis and optimization tools [40].
- The team at IT4Innovations which is mainly involved in performance evaluation and proof of concept services, as well as the new co-design data repository. It will also participate in POP dissemination and training activities

4 Collaboration on the security domain

PRACE has been collaborating on security with other e-Infrastructures from its beginning. Those collaborations are based on various binding levels. Whereas the collaboration with the Human Brain Project (see subsection 2.5) is on a very strong level, based on the shared use of the PRACE MDVPN network, collaborations with other projects are sometimes on an informational base only.

4.1 Collaboration with WISE

The main binding activity on the security perspective is with the WISE community [19] (WISE Information Security for collaborating E-infrastructures), which was formally established in late 2015. Since then, several EU projects, including XSEDE from the US supported this activity by sending security staff members as contributors to this community. Often those people took over leadership as chairs of some of the existing working groups of WISE. Within WISE, representatives from EGI, EUDAT CDI, EOSC-hub, HBP, PRACE, GÉANT, NRENS, and also XSEDE exchange ideas concerning security related aspects. The aim being providing a reliable global framework where security experts can share information on topics such as risk management, experience on certification processes and threat intelligence.

PRACE has provided the co-chair for the “Security in Big and Open Data –Working Group” (SBOD-WG) and is currently the co-chair of the newly founded “Security challenges for high-throughput data transfers – Working Group”. As co-author of the “Security for Collaborating Infrastructures Trust Framework - Versions 1 and 2” [41] [42], PRACE expanded its visibility as one of the leading partners in the security area continuously. The relevance of this document can be valued, when considering that several EU projects/infrastructures have endorsed this document already ([EGI](#), [EUDAT](#), [GÉANT](#), [GridPP](#), [HBP](#), [MYREN](#), [PRACE](#), [SURF](#), [WLCG](#), [XSEDE](#)). Current activities on standardisation are focused on the provision of a standard “WISE Baseline Acceptable Use Policy” which again will presumably be adopted by most of these infrastructures.

Since WISE does not have its own resources, its activities are based on in-kind contributions from the supporting infrastructures.

4.2 Federated Authentication and Authorisation

The Authentication and Authorisation for Research and Collaboration (AARC) [18] initiative was first launched in May 2015 for a two-year period to address the increased need for federated access and authentication and authorisation mechanisms through research and electronic infrastructures. The second phase (AARC2) focused on the development and management of an integrated interdisciplinary authentication and authorisation framework, building on existing authentication and authorisation infrastructures (AAI). This collaboration is two folded. Since AARC2 has an activity to find personnel for collaboration with other projects, e.g. PRACE, PRACE in turn provided capabilities for its partners to work on the AEGIS group “AARC Engagement Group for Infrastructures”. Here, the uptake of AARC2 developed strategies, policies and blueprints are discussed. On the other hand, AARC developments can be based on requirements coming up from the engagement group members, i.e. the participating infrastructures. Outcome of these activities

should become visible in PRACE-6IP, where the idea is to let GÉANT engage in building the PRACE AAI on AARC blueprint architecture (AARC-BPA [43]) design. Prerequisites to this will be to install a PRACE Identity Provider (IdP) and PRACE Attribute server. First discussions on GDPR issues and related security risks have just started.

Having this in mind, it has already been mentioned in the deliverable D6.5 that for authentication and authorisation services it is valuable for PRACE to use services provided by the Identity Provider (IdP) federations which exist for educational organisations in several countries and which are also federated in eduGAIN. This would allow users to use these credentials for both, PRACE access and services provided by other infrastructures. However, it will be essential to trust those federations. Discussions on those issues have already started and are the main prerequisites for adopting the AARC BPA.

Generally, the interaction of PRACE with other infrastructures on security issues is not only based on de jure formal channels, but also on de facto secondary channels. This means that some people may work for several security projects or infrastructures. For example, the PRACE CSIRT leader serves also as Deputy EUDAT IT Security Officer and is also involved in the HBP Security Policy development team.

4.3 Trusted Introducer service

As third collaboration in the security area, PRACE is participating in the GÉANT TF-CSIRT Trusted Introducer service [44]. Indeed, for several years now, PRACE has been leading an internal security team (PRACE CSIRT) to manage security incidents within the infrastructure. As PRACE is a European infrastructure, it is important to be able to collaborate with CSIRTs across Europe and beyond through global collaborations.

The Trusted Introducer services differentiate between four categories:

- Teams are:
 - listed, which provides basic information about the team itself as well as shows endorsement of the team by the TI community;
 - accredited, which ensures a defined level of the best practices and acceptance of the established TI policies for such teams;
 - certified, if they have been accredited before and prove a confirmed level of maturity as defined by the TI SIM framework.
- Security experts can participate as TI Associates.

During PRACE-5IP, PRACE-CSIRT obtained the status “Listed” as we can see in the Figure . A lot of work has been done during PRACE-5IP to lay the basis for getting “Accredited” status in the TI service. The main advantage of this accreditation is that it allows collaboration with other CSIRTs and access to information about incidents and vulnerabilities not otherwise available. It is expected that, presumably at the beginning of PRACE-6IP, this accreditation can be reached.



PRACE CSIRT

PRACE Computer Security Incident Response Team

Listed
since 19 Apr 2018

This information is provided without guarantee or pro-active maintenance.

Fields describing the team

Team Details

| | | |
|---|----------------------------------|--------------------|
| Official Name PRACE Computer Security Incident Response Team | Short Name PRACE CSIRT | Country *Europe |
| Established 02 Apr 2012 | Host Organisation PRACE-RI.EU | |

Constituency

| | |
|---|---|
| Constituency Type Research & Education | Country of Constituency - |
| ASNs, Domains, IP ranges --- | Description PRACE-CSIRT is the Computer Security Incident Response Team compounding the security experts of the PRACE partner institutions. PRACE is established as an international not-for-profit association (aisbl) with its seat in Brussels. It is named 'Partnership for Advanced Computing in Europe AISBL' with 25 member countries. The computer systems and their operations accessible through PRACE provided by T0 and T1 partners are globally distributed through Europe. The systems are accessible via GEANT, the respective NRENS and the local partner networks. A PRACE-VPN connects the systems via a trusted dedicated network not accessible from outside. In case of problems with a particular IP the usual way to contact the PRACE-CSIRT would be via RIPE-DB, NREN and partner contact information, which in turn contacts PRACE-CSIRT directly. |

Team Contact Information

| | | |
|--|--------------------------------------|--|
| Main Number +49 2461 61 4772 | Emergency Number +49 2461 61 6440 | Fax Number +49 2461 61 6656 |
| Email csirt@prace-ri.eu | Other contact - | Postal Address PRACE Office attn: CSIRT Rue du Trone 98 1050 Ixelles |
| Business Hours 08:00 - 17:00 Monday to Friday | Timezone GMT+01 / GMT+02 | |

Classification

| | |
|--|----------------------------|
| Current State Listed since 19 Apr 2018 | Entry Date 29 Mar 2018 |
| Date of Accreditation - | Date of Certification - |

History

| Date | Description |
|-------------|----------------------------------|
| 19 Apr 2018 | PRACE CSIRT is now a listed team |

Figure 5: Team info of the PRACE CSIRT [45]

5 Conclusion

The aim of the general framework of the H2020 programme (EOSC/EDI) is to provide Europe with a world-class online research infrastructure, including reliable and ubiquitous services for networking and computing while providing users with a comprehensive approach integrating all these major research e-Infrastructures. To this end, the most important factor is to understand the interaction between the different service providers (e.g. EUDAT, EGI, GÉANT, etc.) and to contribute to a fruitful cooperation between them.

In this deliverable, we have reported on PRACE-5IP effort spent in this direction, in particular by strengthening our relationships with e-Infrastructures and CoEs, but also by establishing new ones with other projects. We reported on the collaborations that have been conducted during the PRACE-5IP project based on user-driven pilot projects.

We presented in detail the collaboration between PRACE and EUDAT, which has resulted in enriching the offer towards the users, enabling them to utilize more efficient and complete workflows. To demonstrate this, we reported on the pilots that we ran highlighting the benefits coming from a common HPC and data management. Indeed, these projects implement the services built as part of these collaborations and have allowed us to validate the process or highlight the deficiencies of the services provided. Collaborations with this type of project are essential to raise the level of scientific excellence; the requirements resulting from feedback indicate the directions to be taken towards improving or creating services and it is a real virtuous circle for scientific research.

We described the PRACE's collaboration on the operational domain, in relation to the overall HPC European Eco-System, with a particular mention to the CoEs. We have unequivocally understood that they contribute to strengthening European leadership in the field of HPC applications and cover important areas. Indeed, they provide services such as: the development, optimization and scaling of application codes from high-performance computing to peta and exascale computing; quality assurance; consulting to companies and SMEs; research in HPC applications, etc. Therefore, we investigated the potential interaction between the PRACE's operational services and CoEs and reported about the ongoing collaborations.

We discussed the activity on security, which is at the moment a major concern for research infrastructures and for all the actors surrounding it and which needs to be treated as a common cross-activity between all the European e-Infrastructures in a persistent way. Collaboration with GÉANT, AARC and WISE are examples of work made in this direction, as well as PRACE and EUDAT, which have long collaborated in joint security training and workshops for site security officers, sharing best practices and experiences with the joint Security for Collaborating Infrastructures (SCI) Trust Framework.

The HPC European Eco-System is a complex landscape in which PRACE is working actively to collaborate with other actors. This is crucial to provide the users with a modern and comprehensive service that will enable them to better satisfy their computational needs. In this deliverable, we reported the activity of the PRACE-5IP project from start to finish.

For the 6th implementation of PRACE, we look forward to make these collaborations more sustainable and establish new ones. To this end, we will foster relationships with the main European service providers, continuing to focus on improving all security processes by implementing them in different services and projects. We will develop links with CERN, through projects such as EAGIS by defining needs and follow-up of intensive data experiments such as the LHC and SKA. Indeed, many areas for improving the PRACE service catalogue have been identified, in particular the ease of use and flexibility of operating models. The integration of new services into the PRACE catalogue will be ensured with the arrival of new CoEs (EXCELLERAT, ChEESE [46], HiDALGO), through the renewal of most of the existing ones but also through the FocusCoEs initiative [47]. The broad outline of the future directions of this task will be further refined at the EuroHPC Summit week in Poznan and the PRACE 6IP kick-off meeting in Bratislava in May 2019.