



PARTNERSHIP FOR ADVANCED COMPUTING IN EUROPE

PRACE-5IP Overview





PARTNERSHIP FOR ADVANCED COMPUTING IN EUROPE

Evolution of PRACE

2015



PP 1IP 2IP 3IP 4IP Established RRI

Implementation Phase



PARTNERSHIP FOR ADVANCED COMPUTING IN EUROPE

PRACE Projects



PRACE RI

Belgian international non for profit association: PRACE aisbl

Hosting Members (Initial Phase 2010-2015):



Governance : MB and TB

Governance : Council and BoD



Result:
Deliverables



Implementation &
development
PRACE aisbl

Result:
Provide world-class HPC systems
for world-class science
HPC ecosystem



PRACE-5IP Facts

- Project duration: January 2017 – April 2019
- Consortium: 26 partners from 25 countries including PRACE aisbl (BE)
- Budget: 15,9 Mio €
- EC contribution: 15 Mio €
- Planned Budget usage:
 - Workforce: 1496,4 PM = 124,7 FTE for 28 months
 - Dissemination and Training: 1.457 k€ (1.457 k€ EC contribution)
 - Network connectivity 469 k€ (235 k€ EC contribution)

Mandatory Acknowledgment:

This project receives funding from the EU's Horizon 2020 research and innovation programme (2014-2020) under grant agreement no. **730913**.



PRACE-5IP Structure & Objectives

- WP 1: Management of the Contract
- WP 2: Organisational Concept of the RI
- WP 3: Communication and Dissemination
- WP 4: Training
- WP 5: HPC Commissioning and Prototyping
- **WP 6: Operational Services for the HPC Eco-system**
 - Operate and monitor common operational services for the use of the European HPC infrastructure
 - **Analyse new prototypal innovative services**
 - Link with other e-infrastructures and CoEs
- WP 7: Applications Enabling and Support



Design and Development of new services prototypes

Continue the evaluation of some of the pilots tested in 4IP Task 6.2

- Provision of urgent computing services
- **Link with large-scale scientific instruments**
- Smart post processing tools including in situ visualisation
- Provision of repositories for European open source scientific libraries and applications

New Pilots: analyse the potential of new services and develop prototype solutions

- Evaluate lightweight virtualisation technology, such as Docker
 - creation of specialised environments where, to prototype new applications and services
 - distribution of PRACE applications as an alternative to the standard module-load approach
- **Evaluate new prototypes for Data Analytics services**
 - Apache Spark, an open source computing framework for processing large datasets
 - Its capability to cache datasets in memory makes the framework appropriate for large data analysis applications
 - test the technology on a HPC system via a pilot scientific use case and evaluate the scalability benefits
- Investigate the possible adoption in a next phase as production services



Data Analytics Service description

“Evaluation of new prototypes for Data Analytics services, such as Apache Spark, an open source computing framework for processing large datasets. Its capability to cache datasets in memory makes the framework appropriate for large data analysis applications, especially on systems with large memory capacity. The purpose of this activity is to test the technology on a HPC system via a pilot scientific use case and to evaluate the scalability benefits. For these prototypes, the possible adoption in a next phase as production services will be investigated, analysing for them the aspects of service certification. Recommendations for adoption as production services will be made to the PRACE aisbl. “



High level activities

- Build the team
- Gather requirements to validate/feed task activities
 - PRACE projects?
- Select relevant frameworks to anticipate user needs
 - Data Analytics
 - Machine/Deep Learning
- Evaluate performance by piloting concrete use case
- Optimize/tune the framework on PRACE systems
- Package a service/module / container



Data Analytics

- Possible evaluations with Apache Spark (open source computing framework for processing large datasets)
 - Apache Spark + RDMA
 - <https://github.com/intel-analytics/BigDL>
 - <http://hibd.cse.ohio-state.edu/>
 - User level I/O architecture with Crail
 - Crail is an open source user-level I/O architecture for the Apache data processing ecosystem
<http://www.crail.io/>



Deep Learning

- Select relevant framework
 - TensorFlow/Caffe/Theano/Intel-neon
- Evaluate/test backend engines performance
 - Compute (low level libraries for many cores and GPUs)
 - MKL-DNN vs. cuDNN
 - MKL-BLAS vs. cuBLAS
 - Communication
 - Intra node: NCCL (nVIDIA Collective Communication Library)
 - Inter node
 - gRPC (Google Remote Procedure Call)
 - INTEL-MLSL (Machine Learning Scaling Library)