



Overall Goals and Opportunities for Collaboration

High Performance Computing Collaboration Agreement - Technical Coordinators

Florian Berberich, Vincenzo Capone, Miles Deegan, Chiara Ferrari, [Maria Girone](#)

Kick-off Workshop on HPC, 29.09.2020

Motivation

- The collaboration agreement has been signed by GÉANT, PRACE, CERN and SKAO on 22.07.2020
- Engages at the community level and today we have good community participation
 - It comes at an opportune time as HPC centers are preparing build to exascale computing resources
- Data intensive sciences like HL-LHC and SKA will be exascale sciences
 - In the presence exponentially growing needs and essentially flat budgets the effective use of HPC sites will be critical
- This agreement is the result of discussions between the four parties that began in 2018

AENEAS All-Hands Bologna Oct 8-10, 2018 - <https://indico.astron.nl/conferenceDisplay.py?confId=172>
CERN/PRACE Workshop Oct 22, 2018 - <https://indico.cern.ch/event/760705>

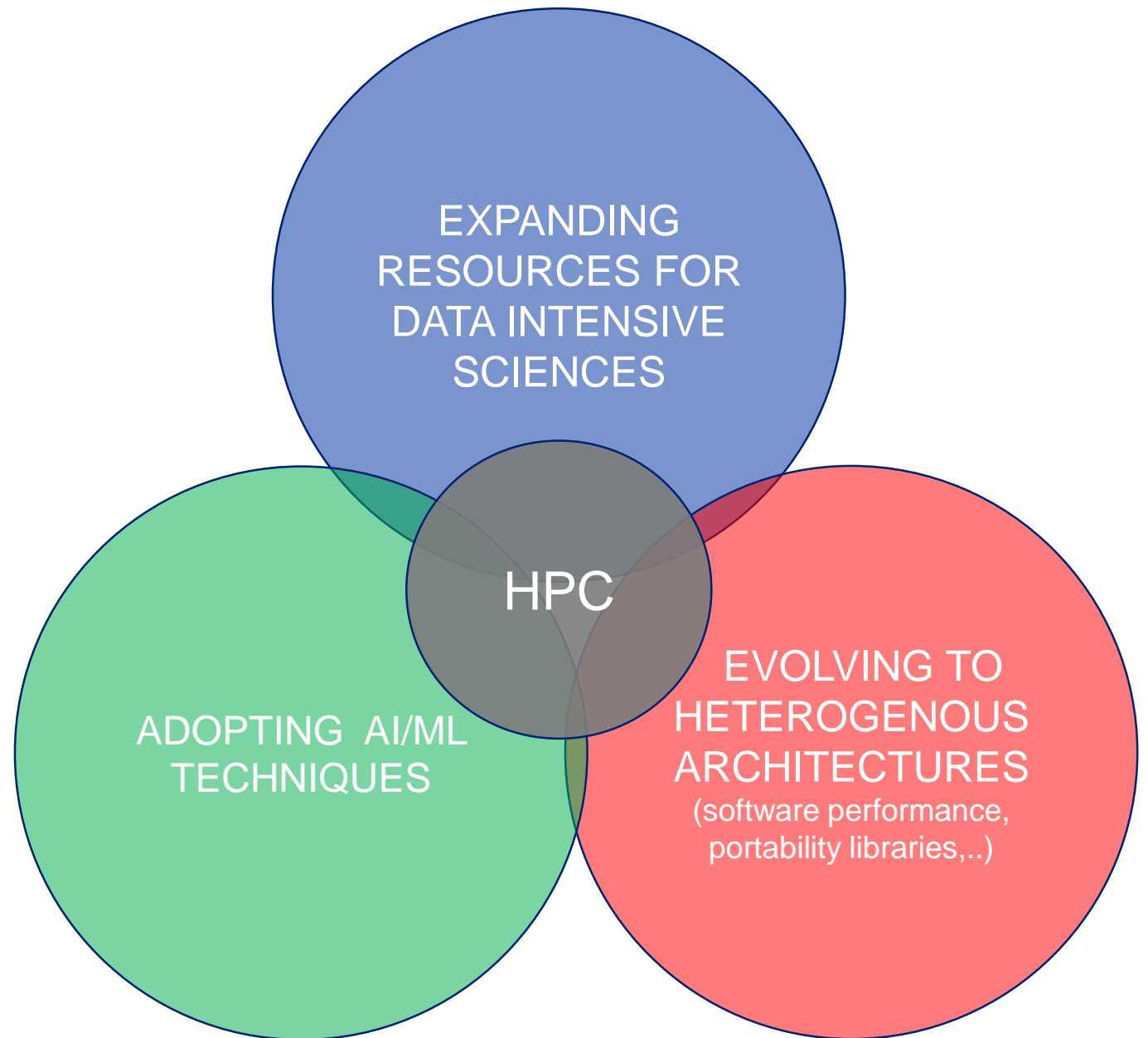


Eckhard Elsen (top left), Director for Research and Computing at CERN; Philip Diamond (top right), SKA Director-General; Erik Huizer (bottom left), Chief Executive Officer of GÉANT; and Philippe Lavocat (bottom right), PRACE Council Vice-Chair, signed the agreement for the new collaboration.

Synergies

HPC falls at the intersection of several important R&D areas

- Engagement with the HPC community can be a catalyst for progress
- This collaboration is a forum and foundation for development work
- We will rely on synergies with other projects (ESCAPE)



Workshop Goals

- **This is a brainstorming and team-forming event**

- Active participation is key
- There is a Google Doc for Live Notes

https://docs.google.com/document/d/1yss1gDOtsH_-O_vL-saKnt1cwR3I_3Lr1nYomwKe4C0/edit?usp=sharing

- **We are not operating in a vacuum**

- Challenges in integrating HPC for data intensive science have been defined by the experiments
- There are existing projects in some areas
- We will build on existing activities as we define common directions among the four organizations

HPC resources integration at LHCb

https://docs.google.com/document/d/1Wu_jvIzQASdsfxEi3Twe2n_QQWC38wsZYvDv-sReU14/edit?usp=shari

HPC resources at CMS

https://cds.cern.ch/record/2707936/files/NOTE2020_002.pdf?version=1

HEP Software Foundation Community White Paper

<https://hepsoftwarefoundation.org/cwp-whitepapers.htm>

Common Challenges for HPC Integration into LHC Computing

https://zenodo.org/record/3647548#.X28uIS2z0_W

- **Four areas of work have been identified as foundational. Progress will be evaluated by a series of challenges and demonstrators**

- Training and center of expertise
- Authentication and Authorization
- Benchmarking
- Data Access

- **The goal of this meeting is to begin to form programs of work**

- Define metrics to measure progress

The technical program (1)

Training and Center of Expertise (Convener F. Berberich, PRACE)

- Our goal is to leverage expertise from the collaboration for training and support
 - PRACE for software and scale for HPC
 - GÉANT for networking
 - CERN and SKA for data intensive science and data management
- We would like to capitalize on the technical expertise in the collaboration and make training plans to support the community

The technical program (2)

Authorization and Authentication (Convener Rosie Bolton, SKA)

- HPC sites often have stricter cybersecurity due to their size and capabilities
- Data intensive sciences are larger organizations than many HPC communities
- We will discuss the needs and goals on each side to arrive at a plan for securely supporting data intensive science workflows on HPC
 - There is the potential to align with broader on-going AAI modernization efforts

The technical program (3)

Benchmarking (Convener M. Girone, CERN)

- Needed for evaluating progress towards integrating HPC
- There is an existing prototype based on on-going HEP of a containerized standalone benchmark suite
 - Intended to be lightweight and extensible to additional scientific applications
 - Prototype contains several HEP applications and has been run at HPC scale

The technical program (4)

Data Access (Convener V. Capone, GÉANT)

- Data access is a critical challenge for data intensive science
 - A data access demonstrator is needed to show that HPC sites can be effectively used
 - The ability to move petabytes
 - The ability to maintain processor efficiency with high data access rate and to export results
 - Demonstrate the ability to read locally and stream data

Conclusions

- **The Collaboration on High Performance Computing is a unique opportunity to make progress on identified challenges in using HPC facilities**
- **Software performance for heterogeneous architectures, benchmarking and accounting, data access, authentication and authorization are areas of significant common interest**
- **Our goal is to kick-off common investigations around these areas and share requirements but also the progress that we are making as communities**